<u>What is claimed is:</u>

1.   A rule-based speech synthesis device which synthesizes arbitrary speech by selecting and concatenating previously stored speech synthesis units and controlling prosodic information, comprising:

a phoneme duration time setting means which estimates and controls the closing interval length of a phoneme having a closing interval, independently of the vowel length and consonant length.

2.   The rule-based speech synthesis device according to claim 1, wherein said phoneme duration time setting means comprises:

a phoneme type judgement means that judges the type of a phoneme with respect to the input phoneme symbol sequence,

a vowel length determining means comprising a vowel length estimation means and a vowel length learning means,

a consonant length determining means comprising a consonant length estimation means and a consonant length learning means, and

a closing length determining means comprising a closing length estimation means and a closing length learning means;

and wherein said phoneme type judgement means operates said vowel length estimation means or consonant length estimation means depending on whether the phoneme in question is a vowel or a consonant, and if it is judged to be a consonant, it judges whether or not it anteriorly has a closing interval and if it

anteriorly has a closing interval then it operates a closing length estimation means.

3. The rule-based speech synthesis device according to claim 2, wherein:

said closing length determining means further comprises a closing length classification means;

said closing length classification means performs classification operations whereby it obtains a frequency distribution of closing lengths from learning data, classifies the closing lengths into a first group based on said frequency distribution and classifies the phoneme in question into a second group based on the first group;

said closing length learning means performs learning operations whereby it is learned with each member of the said second group and outputs weighting coefficients - which are necessary for estimation of phoneme duration times - to the closing length estimation means; and

said closing length estimation means judges the name of the phoneme in question from an input phoneme symbol sequence, judges and selects said second group from said phoneme name, selects weighting coefficients inherent to said group, performs operations to estimate the closing length using said weighting coefficients, and outputs the value of the estimated closing length.

4. The rule-based speech synthesis device according to claim 2, wherein:

said vowel length determining means further comprises a vowel length classification means;

said vowel length classification means performs classification operations whereby it obtains a frequency distribution of vowel lengths from learning data, classifies the vowel lengths into a first group based on said frequency distribution and classifies the phoneme in question into a second group based on the first group;

said vowel length learning means performs learning operations whereby it is learned with each member of the said second group and outputs weighting coefficients - which are necessary for estimation of phoneme duration times - to the vowel length estimation means; and

said vowel length estimation means judges the name of the phoneme in question from an input phoneme symbol sequence, judges and selects said second group from said phoneme name, selects weighting coefficients inherent to said group, performs operations to estimate the vowel length using said weighting coefficients, and outputs the value of the estimated vowel length.

5.   The rule-based speech synthesis device according to claim 2, wherein:

said consonant length determining means further comprises a consonant length classification means;

said consonant length classification means performs classification operations whereby it obtains a frequency distribution of consonant lengths from learning data, classifies

the consonant lengths into a first group based on said frequency
distribution and classifies the phoneme in question into a
second group based on the first group;

said consonant length learning means performs learning
operations whereby it is learned with each member of the said
second group and outputs weighting coefficients - which are
necessary for estimation of phoneme duration times - to the
consonant length estimation means; and

said consonant length estimation means judges the name of
the phoneme in question from an input phoneme symbol sequence,
judges and selects said second group from said phoneme name,
selects weighting coefficients inherent to said group, performs
operations to estimate the consonant length using said weighting
coefficients, and outputs the value of the estimated consonant
length.

6.   The rule-based speech synthesis device according to
claim 3, wherein:

said closing length learning means is composed of a first
factor extraction means which extracts and quantizes factors
comprising the phoneme in question, the phoneme environment
consisting of the two phonemes before and after the phoneme in
question, the phoneme position, the part of speech and the like,
a first prior de-voicing judgement means which judges whether or
not the previous phoneme is de-voiced based on the learning
data, and a model learning means which produces weighting
coefficients for each factor in each of said classified second
groups;

and wherein said closing length estimation means is composed of a second factor extraction means which extracts and quantizes factors comprising the phoneme in question, the phoneme environment consisting of the two phonemes before and after the phoneme in question, the phoneme position, the part of speech and the like, a second prior de-voicing judgement means which judges whether or not the phoneme in question is to be de-voiced based on prescribed de-voicing rules, and a model estimation means which judges said second group from the phoneme in question and estimates the closing length by referring to the weighting coefficients output from said model learning means for each group.